Bioorganic & Medicinal Chemistry 17 (2009) 7635–7642

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Peptide inhibitors of HIV-1 integrase: From mechanistic studies to improved lead compounds

Michal Maes [a], Aviad Levin [b], Zvi Hayouka [a], Deborah E. Shalev [c], Abraham Loyter [b], Assaf Friedler [a,*]

[a] Institute of Chemistry, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel
[b] Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel
[c] The Wolfson Centre for Applied Structural Biology, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel

ARTICLE INFO

Article history:
Received 18 May 2009
Revised 22 September 2009
Accepted 30 September 2009
Available online 4 October 2009

Keywords:
HIV-1
Integrase
Peptides
Inhibitors

ABSTRACT

The HIV-1 integrase enzyme (IN) catalyzes integration of viral DNA into the host genome. We previously developed peptides that inhibit IN in vitro and HIV-1 replication in cells. Here we present the design, synthesis and evaluation of several derivatives of one of these inhibitory peptides, the 20-mer IN1. The peptide corresponding to the N-terminal half of IN1 (IN1 1–10) was easier to synthesize and much more soluble than the 20-mer IN1. IN1 1–10 bound IN with improved affinity and inhibited IN activity as well as HIV replication and integration in infected cells. While IN1 bound the IN tetramer, its shorter derivatives bound dimeric IN. Mapping the peptide binding sites in IN provided a model that explains this difference. We conclude that IN1 1–10 is an improved lead compound for further development of IN inhibitors.

1. Introduction

The human immunodeficiency virus type 1 (HIV-1) integrase (IN) catalyzes integration of the reverse-transcribed viral DNA into the host cell genome. This is an essential step in the viral replication cycle. IN has no mammalian homologues and inhibiting IN is a major target for specific anti-retroviral drugs.[1–4] IN mediates integration in two steps:[5] (I) 3′-end processing, where IN removes a pGT dinucleotide from the 3′ end of each strand of the linear viral DNA.[5] This step occurs after reverse transcription in the cytoplasm and is carried out by two IN dimers that bind the DNA in its two long terminal repeats (LTR) termini;[6] (II) strand transfer: integrating the viral DNA into the target host DNA, which takes place in the nucleus. Tetrameric IN is required for strand transfer activity.[7] IN is in equilibrium between its dimeric and tetrameric oligomerization states, both of which are required for the various stages of the integration reaction.[8–10] Structurally, IN comprises three domains: an N-terminal zinc-binding domain (residues 1–55),[11] a catalytic core domain (CCD; residues 50–212),[12] and a C-terminal DNA binding domain (residues 220–270).[13] A catalytic triad composed of residues D64, D116 and E152 in the CCD is responsible for the enzymatic activity of IN.[12]

Several peptidic IN inhibitors have been reported in the literature.[14–16] We have developed several peptides that inhibit IN activity in vitro and in cells as well as HIV-1 replication in cells.[17–20] All these inhibitory peptides bound preferentially to the IN tetramer compared to the dimer and stabilized the IN tetramer.[17–19] We have termed these peptides ‘shiftides’, since by stabilizing the IN tetramer they may shift the IN oligomerization equilibrium.[6] It is still unclear how our peptidic IN inhibitors induce IN tetramerization, and whether this activity is sufficient and/or necessary for inhibiting IN.

The IN1 peptide, which is the focus of this study, was selected in our laboratories from a combinatorial library using the yeast two-hybrid system.[17] IN1 bound the IN tetramer with a K_d of 8.7 μM.[17] It is a linear 20-residue peptide and thus its pharmacological properties in general and stability in particular are not optimal.[21] In addition, it is very hydrophobic, making it difficult to achieve high synthetic yields due to insolubility. Here, shortened derivatives of IN1 were developed to improve IN1 as an anti-HIV lead compound as well as to study its mechanism of IN binding. Our results reveal that while the 20-residue IN1 binds preferentially to the tetrameric IN, the 10-residue peptides bind mainly to the dimer. We present a model that may explain this selectivity based on analyzing the peptide binding sites in IN. The 10-residue IN1 derivative IN1 1–10 is more active, more soluble and easier to synthesize than the parent IN1 peptide and is an improved lead compound for further developing IN inhibitors.

2. Results

2.1. Design and synthesis of modified IN1-derived peptides

Several shortened and modified derivatives of the IN1 peptide were designed and synthesized (Table 1a). The original 20-mer

* Corresponding author. Tel.: +972 2 6585746; fax: +972 2 6585345.
E-mail address: assaf@chem.ch.huji.ac.il (A. Friedler).

0968-0896/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2009.09.053

Table 1
Peptides used in this study

Peptide name	Sequence[a]
(a) Sequences	
IN1	WQCLTLTHRGFVLLTITVLR
IN1 1–10	WQCLTLTHRG
IN1 11–20	WFVLLTITVLR
IN1 1–10 C3S	WQSLTLTHRG
IN1 16–20	WITVLR

Peptide	Calculated MW	Observed MW (MALDI-TOF)	Purity (Analytical HPLC) (%)
(b) Molecular weights and degree of purity[b]			
IN1	2370	2369	93
IN1 1–10	1214	1213	99
IN1 11–20	1360	1361	90
IN1 1–10 C3S	1198	1197	98
IN1 16–20	786	785	93

	HN	Hα	Hβ	Others
(c) NMR assignment (δ_{1H} ppm)[c]				
IN1				
Trp1	—	4.27	3.35	Aromatic 10.21, 7.52, 7.48, 7.25, 7.16, 7.08
Gln2	8.36	4.32	2.20	Hγ 1.93, 1.84
Cys3	8.43	4.30	2.80	
Leu4	8.54	4.40	1.56	Aliphatic 0.86, 0.79
Thr5	8.21	4.31	4.12	Hγ 1.14
Leu6	8.29	4.35	1.56	Aliphatic 0.84, 0.82
Thr7	8.02	4.28	4.13	Hγ 1.11
His8	8.15	4.62	3.01, 2.97	Aromatic 8.52, 7.22
Arg9	8.37	4.28	1.79, 1.67	Hγ 1.55, Hδ 3.14
Gly10	8.34	3.91, 3.82		
Phe11	8.53	4.64	3.23, 3.12	Aromatic 7.22, 6.95
Val12	8.10	4.02	1.91	Hγ 0.82
Leu13	8.27	4.29	1.53	Aliphatic 0.90, 0.84
Leu14	8.30	4.40	1.54	Aliphatic 0.88, 0.82
Thr15	8.12	4.31	4.11	Hγ 1.13
Ile16	8.20	4.21	1.81	Aliphatic 1.58, 1.41, 0.88
Thr17	8.26	4.23	4.07	Hγ 1.13
Val18	8.17	4.05	2.00	Hγ 0.88
Leu19	8.34	4.33	1.58, 1.55	Aliphatic 0.88, 0.82
Arg20	8.31	4.26	1.82, 1.72	Hγ 1.61, 1.57; Hδ 3.15
IN1 1–10				
Trp1	—	4.26	3.35	Aromatic 10.27, 7.54, 7.48, 7.28, 7.21, 7.11
Gln2	8.32	4.27	2.21	Hγ 1.90
Cys3	8.41	4.28	2.83	
Leu4	8.53	4.39	1.60	Aliphatic 0.85
Thr5	8.12	4.28	4.15	Hγ 1.16
Leu6	8.26	4.38	1.58	Aliphatic 0.87
Thr7	8.04	4.25	4.10	Hγ 1.13
His8	8.48	4.71	3.18	Aromatic 8.53, 7.16
Arg9	8.48	4.31	1.78	Hγ 1.61; Hδ 3.18
Gly10	8.49	3.89		
IN1 11–20				
Trp	—	ND	ND	Aromatic 11.00, 7.36, 7.28, 7.06, 6.82, 5.63, 5.62
Phe11	7.89	3.95	3.24, 3.00	Aromatic 7.17, 7.09, 7.00
Val	7.67	4.14	1.75	Hγ 0.82
Val	7.62	4.04	1.99	Hγ 0.83
Leu	8.15	4.05	205	Aliphatic 0.87
Leu	7.95	4.19	1.61, 1.46	Aliphatic 0.85
Leu	7.92	4.21	1.64, 1.49	Aliphatic 0.82
Thr	7.89	4.24	1.48	
Thr	7.68	4.26	1.50	
Ile16	7.79	4.14	1.50, 1.67	Aliphatic 1.02
Arg20	7.83	4.20	1.70	Hγ 1.51; Hδ 3.06, Hε 7.44
IN1 1–10 C3S				
Trp1	—	4.27	3.35	Aromatic 10.17, 7.47, 7.22, 7.06, 6.82
Gln2	8.31	4.27	2.22	Hγ 1.96, 1.85
Ser3	8.32	4.24	3.79	
Leu4	8.44	4.36	1.59	Aliphatic 0.85
Thr5	8.15	4.26	4.11	Hγ 1.15
Leu6	8.22	4.36	1.58	Aliphatic 0.86
Thr7	8.01	4.23	4.10	Hγ 1.12
His8	8.46	4.67	3.24, 3.14	Aromatic 8.56, 7.24
Arg9	8.47	4.30	1.81, 1.73	Hγ 1.59; Hδ 3.14; Hε7.14
Gly10	8.49	3.88		
IN1 16–20				
Trp	—	4.35	3.31, 3.39	Aromatic 10.21, 7.60, 7.54, 7.49, 7.24, 7.15, 7.08
Ile16	8.32	4.22	1.75	Aliphatic 1.38, 1.07, 0.82
Thr17	8.27	4.28	4.06	Hγ 1.16
Val18	8.24	4.05	2.00	Hγ 0.88
Leu19	8.37	4.35	1.57	Aliphatic 0.89, 0.82
Arg20	8.34	4.26	1.81, 1.72	Hγ 1.60; Hδ 3.16; Hε 7.18

[a] Peptides were synthesized and purified as described in Section 4. Trp was added to the N-termini of the peptides in order to allow accurate determination of their concentration by UV spectroscopy.
[b] HPLC traces in Figure S1.
[c] All values for water samples except for IN1 11–20W taken in DMSO-d_6.

peptide was divided into two shorter peptides termed IN1 1–10 and IN1 11–20, which correspond to its N- and C-terminal halves. A shorter peptide derived from the five C-terminal residues of IN1, termed IN1 16–20, was also synthesized. To prevent the formation of disulfide-bridged peptide dimers, we also designed modified peptides where cysteine 3 was replaced by a serine (IN1 C3S and IN1 1–10 C3S).

The sequence of the original IN1 peptide is highly hydrophobic (Table 1). This resulted in low yields during the synthesis of full-length IN1, IN1 11–20 and IN1 C3S. The synthetic problems included severe on-resin aggregation and high insolubility during the purification process. Even incorporating pseudoproline building blocks[22,23] did not improve the yield. However, synthesis on ChemMatrix resin[24–26] (see Section 4) improved yields and gave sufficient amounts of pure peptides. IN1 1–10 was much more soluble and did not show any synthetic problems.

2.2. IN binding to shorter IN1 derivatives shows different characteristics than to the parent IN1 peptide

IN binding to the peptides was determined using fluorescence anisotropy (Fig. 1a). The 10-residue peptides, IN1 1–10, IN1 1–10 C3S and IN1 11–20, bound IN with an affinity of around 1 μM, almost an order of magnitude tighter than the parent IN1 peptide (Fig. 1a and Table 2a). The cysteine to serine replacement did not have an effect on the affinity of the peptides to IN. IN1 16–20 did not show any binding to IN, indicating that these five residues are not sufficient for the interaction of IN with IN1. In contrast to the parent IN1 peptide, which bound the IN tetramer, all 10-residue peptides bound dimeric IN, as reflected by Hill coefficients around 2 (Table 2, Fig. 1a). We used analytical gel filtration to test the effect of peptide binding to IN on its oligomerization equilibrium (Fig. 1b). Free IN (10 μM) eluted from the column at 13.0 ml, corresponding to the size of an IN dimer. In the presence of IN1, the elution volume was 12.6 ml, indicating a shift towards

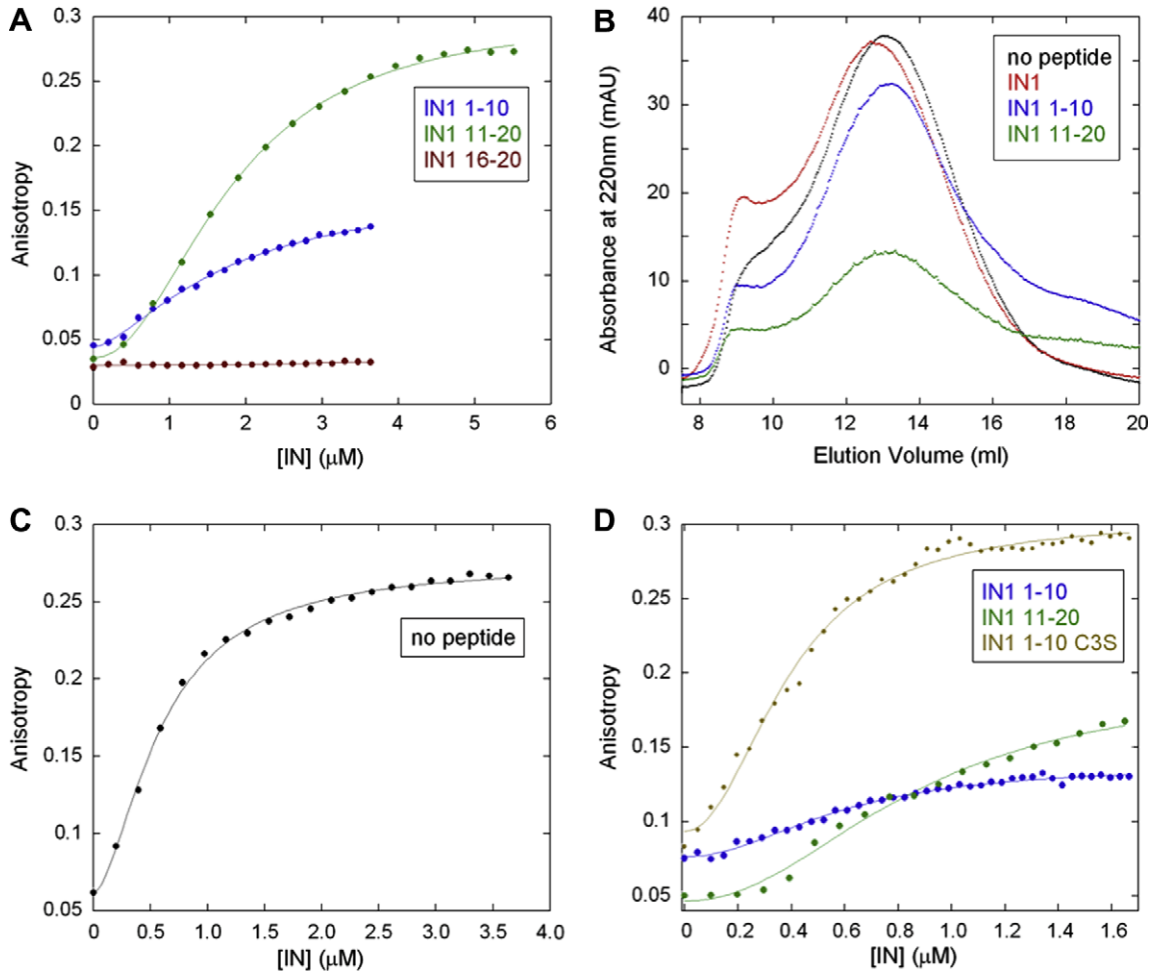

Figure 1. The short IN1-derived peptides bind IN and inhibit its DNA binding in vitro: (A) binding of the short IN1 peptides to IN. The IN protein was titrated into the fluorescein-labeled IN1 1–10 (blue), IN1 11–20 (green) and IN1 16–20 (brown) peptides. Binding affinities and Hill coefficients are shown in Table 2a; (B) effect of peptide binding on the oligomeric state of IN (10 μM); in presence of IN1 (red), a shift in the elution volume of the analytical gel filtration towards a higher oligomeric state can be observed. IN alone (black) or with IN1 1–10 (blue) or IN1 11–20 (green) all elute at the same volume, corresponding to IN dimer; (C and D) the short IN1 peptides inhibit the DNA binding of IN. IN (10 μM) was titrated into fluorescein-labeled viral LTR DNA (10 nM) (C) alone, and (D) after a 20 min incubation on ice with 500 nM of IN1 1–10 (blue), IN1 11–20 (green) or IN1 1–10 C3S (ochre). Binding affinities and Hill coefficients are shown in Table 2b.

Table 2
The IN1 peptides bind IN and inhibit IN-DNA binding[a]

(a) Binding of the IN1-derived peptides to IN		
Peptide	K_d (μM)	Hill coefficient
IN1[17]	8.7	4.5
IN1 1–10	1.7 ± 0.1	1.5 ± 0.1
IN1 11–20	1.8 ± 0.1	2.1 ± 0.1
IN1 1–10 C3S	1.2 ± 0.1	2.0 ± 0.3
IN1 16–20	No detectable binding	
(b) Binding of IN to DNA in presence of the IN1-derived peptides		
Peptide	K_d (nM)	Hill coefficient
No peptide	59 ± 2	1.7 ± 0.1
IN1[17]	300	
IN1 1–10	590 ± 20	2.0 ± 0.1
IN1 11–20	870 ± 90	2.1 ± 0.3
IN1 1–10 C3S	390 ± 10	2.1 ± 0.1

[a] IN binding to (a) the fluorescein-labeled IN1-derived peptides or (b) to fluorescein-labeled DNA in presence and absence of the non-labeled IN1 peptides was determined using fluorescence anisotropy. Binding curves are shown in Figure 1. The results shown are of the data fit described in Section 4.

a dimer–tetramer mixture, consistent with our previous results.[17] In the presence of both short peptides, IN1 1–10 and IN1 11–20, IN remained dimeric and the elution volume did not change, in agreement with the Hill coefficients obtained from our fluorescence anisotropy results. This indicates that IN1, but not its shorter derivatives, stabilizes the IN tetramer.

2.3. The IN1-derived peptides inhibit the DNA binding and catalytic activity of IN in vitro

Fluorescence anisotropy was used to test the effect of the IN1-derived peptides on DNA binding of IN. IN bound fluorescein-labeled LTR DNA with an affinity of 59 nM, and a Hill coefficient around 2 (Fig. 1c, Table 2), in agreement with previous results.[18] IN1 1–10, IN1 1–10 C3S and IN1 11–20 inhibited the IN–DNA interaction by approximately an order of magnitude (Fig. 1d, Table 2). The affinity of IN to DNA in the presence of IN1 11–20 was 870 nM, and in the presence of IN1 1–10 it was 590 nM. Inhibition by IN1 1–10 C3S was slightly weaker. In our previous studies the parent IN1 inhibited the IN–DNA binding affinity at 300 nM.[17] Both IN1 1–10 and IN1 11–20 inhibited DNA binding of IN approximately twofold better than the parent IN1 under the same conditions. In the presence of the peptides, IN still bound the LTR DNA as a dimer, as reflected by the Hill coefficient that remained around 2 (Table 2b).

A quantitative in vitro integration assay was performed to determine whether the peptides inhibit the enzymatic activity of IN (Fig. 2 and Table 3). Experiments were done at an IN:peptide

molar ratio of 1:200 (390 nM IN and 76 μM of the indicated peptides), since it was previously found to be the optimal ratio for inhibition by the full-length IN1.[17] Under these conditions, IN1 inhibited the integration reaction by 77%. IN1 1–10 inhibited the reaction almost to a similar extent, 73%. Inhibition by IN1 11–20 could not be tested at the molar ratio of 1:200 due to the insolubility of the peptide. Thus, inhibition by this peptide was tested at an IN:peptide molar ratio of 1:50, which was the highest peptide concentration that could be achieved. Under these conditions, inhibition by IN1 11–20 was weaker, reaching only 24%. Again, the IN1 1–10 C3S peptide behaved similarly to IN1 1–10, and inhibited the IN activity by 62%.

2.4. The IN1-derived peptides penetrate HeLa cells and inhibit HIV-1 replication in cultured cells by inhibiting viral DNA integration

To test the effect of the peptides on HIV-1 replication in HIV-1 infected cells, we first tested their ability to penetrate cells.[18] Fluorescein-labeled peptides were incubated with cultured HeLa cells and intracellular fluorescence was monitored by confocal microscopy. IN1 1–10, IN1 11–20 (Fig. 3) and IN1 1–10 C3S (data not shown) all penetrated the cytoplasm and the nuclei of the cells.

The MAGI system was used in TZM-bl cells[27] to study the ability of the peptides to inhibit HIV-1 replication in cultured cells. Both the parent IN1 and IN1 1–10 (62.5 μM) inhibited HIV replication by around 70% (Fig. 4a and Table 3). IN1 1–10 C3S inhibited HIV replication by 60% under the same conditions. Inhibition by IN1 11–20 was less pronounced, reaching only a value of 30% at the same peptide concentration. The inhibition of HIV-1 replication by the peptides was concentration dependent.

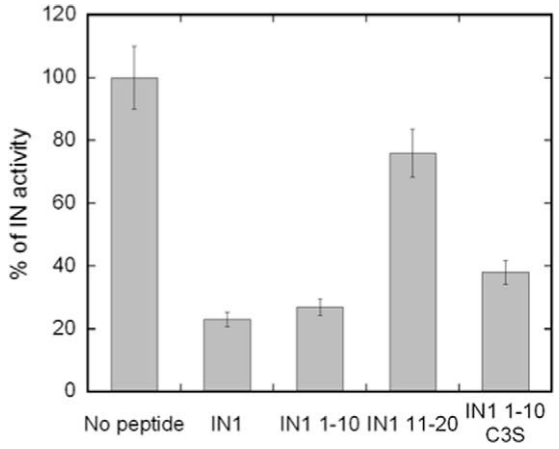

Figure 2. The peptides inhibit the IN catalytic activity in vitro. IN (390 nM) was incubated with each peptide at molar ratios (IN: peptide) of 1:200 (with IN1, IN1 1–10, IN1 1–10 C3S) or 1:50 (with IN1 11–20, see text for details) before performing the IN activity assay as described in Section 4. The first column, marked no peptide, gives the 100% of IN catalytic activity.

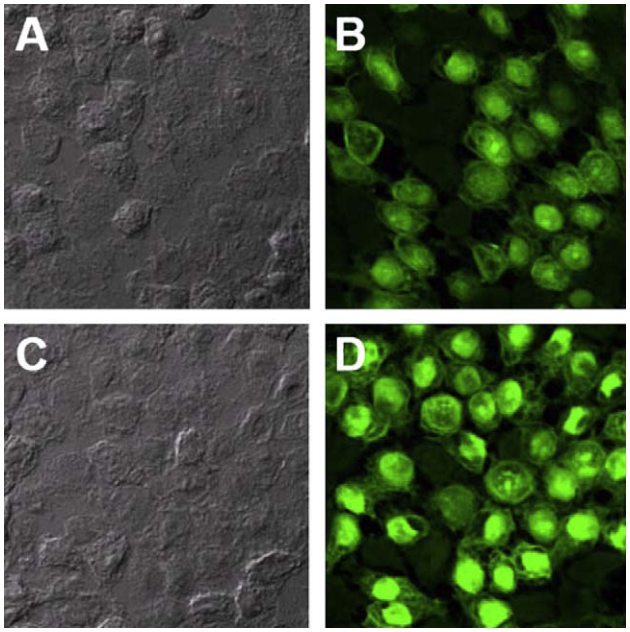

Figure 3. The short IN1-derived peptides penetrate HeLa cells. Fluorescein-labeled peptides were incubated with HeLa cells. The cells were visualized by fluorescent microscopy using a confocal microscope. The fluorescein-labeled peptides localize both in the cytoplasm and the nucleus: (A) IN1 1–10; (B) fluorescein-IN1 1–10; (C) IN1 11–20 and (D) fluorescein-IN1 11–20.

To test whether inhibition of HIV replication by the peptides in cells occurred at the integration step, the integration of the viral DNA into the host cell was measured. Viral DNA was directly estimated using real time PCR 24 h post infection[20] (Fig. 4b and Table 3). IN1 and IN1 1–10 (62.5 μM) inhibited the integration in cells by around 65%. At the same peptide concentrations, IN1 11–20 inhibited only by 23% and inhibition by IN1 1–10 C3S reached 51%.

The total amount of reverse-transcribed viral cDNA in infected cells was estimated to verify that HIV inhibition in cells was not prior to integration. The total amount of viral DNA did not change in presence of the peptides (Fig. 4c), indicating that reverse transcription occurred and the steps prior to integration were not inhibited.

2.5. The IN1-derived peptides bind different sites in IN

The short IN1 derivatives bound IN with different characteristics relative to the parent IN1 peptide. Their affinity was tighter, they bound an IN dimer rather than tetramer, and IN1 1–10 inhibited DNA binding more efficiently (Fig. 1c and d). These differences could be reconciled by showing which sites in IN mediate its interaction with the different IN1 peptides (Table 4 and Fig. 5). ELISA binding studies between the inhibitory peptides and a peptide library covering the full length of the IN protein were performed, similar

Table 3
The IN1 peptides inhibit IN activity and HIV-1 infectivity[a]

Inhibition of:	IN activity (in vitro) (%)	HIV-1 replication (in cultured cells—MAGI) (%)	Integration (in cultured cells—real time PCR) (%)
IN1	77	70	63
IN1 1–10	73	68	64
IN1 11–20	24	29	23
IN1 1–10 C3S	62	60	51

[a] IN activity in vitro (Fig. 2) and HIV infectivity in cells (Fig. 4) were determined as described in Section 4. IN activity in vitro is at an IN:peptide molar ratio of 1:200 (or 1:50 for IN1 11–20, see text for details). Percentage of inhibition in cultured cells is at the highest peptide concentration used (62.5 μM per well) in reference to the initial signal for the HIV replication or in reference to the integration activity in absence of any peptide.

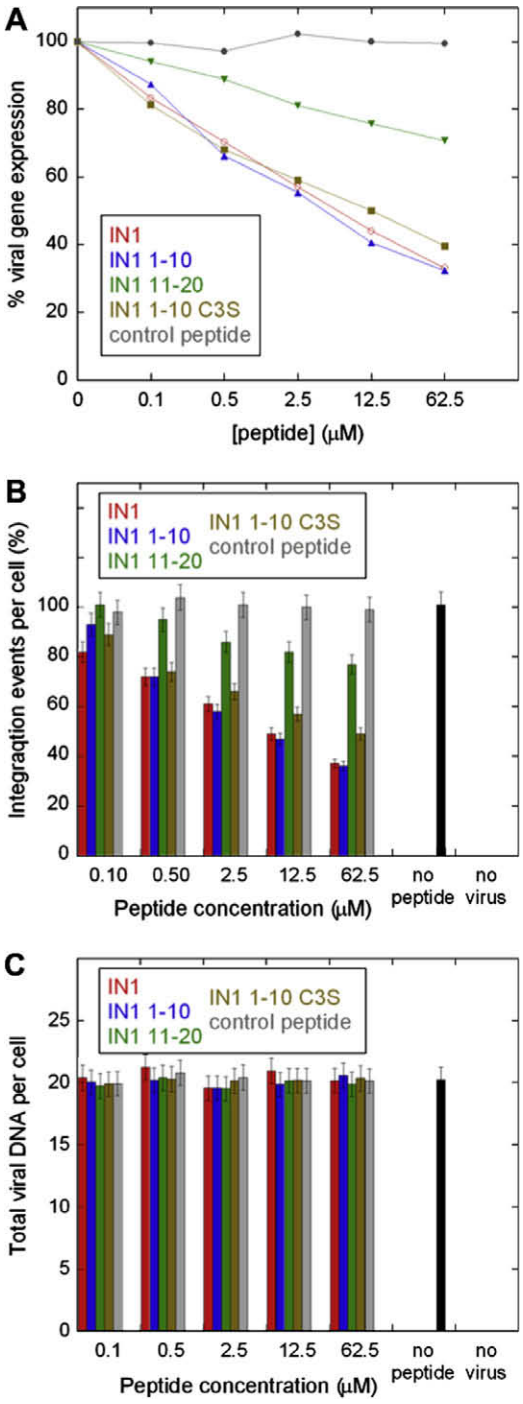

Figure 4. The short IN1-derived peptides inhibit HIV-1 replication in cultured cells: (A) inhibition of HIV infectivity: HeLa MAGI TZM-bl cells were incubated for 2 h pre-infection with different concentrations of the peptides. Blue cells were counted as a measure of the viral early stage of infection. IN1 (red) and IN1 1–10 (blue) show up to 70% inhibition of infectivity, IN1 1–10 C3S shows 60% inhibition, while IN1 11–20 (green) shows only around 30% inhibition at the final concentration of 62.5 μM. An unrelated peptide MZ 6–1 (grey), which is known to penetrate cells, was used as a negative control, showing no inhibition at all; (B and C) SupT1 T-lymphoid cells were incubated with the indicated peptides at the designated concentrations; following HIV-1 infection, the percentage of viral DNA was assessed using real time PCR. In presence of IN1 (red) and IN1 1–10 (blue), up to 65% less integration events can be observed at the final concentration of 62.5 μM. At the same peptide concentration, IN1 1–10 C3S inhibited the integration by 51% and IN1 11–20 only by 23%; (B) quantification of integration events in cells; (C) quantification of total viral DNA; the peptides had no effect on the amount of viral DNA found in the cells at the peptide concentrations used in this experiment. An unrelated peptide,[17] which is known to penetrate cells but without IN inhibitory activity, was used as a negative control, showing no inhibition.

Table 4
The peptide binding regions on IN[a]

Binding region	Interacts with IN1	With IN1 1–10	With IN1 11–20
IN 70–84	V	–	–
IN 94–108	V	V	–
IN 98–112	V	V	–
IN 118–132	V	V	V
IN 242–256	V	–	–
IN 246–260	–	–	V

[a] The regions in IN that bind each peptide were determined using ELISA as described in Section 4. See Figure 5 for the location of these regions on the IN structure.

to those done for IN1.[17] The parent 20-mer IN1 peptide bound five peptides from the IN-derived peptide library, corresponding to four regions in the IN protein (Fig. 5a): residues 70–84, 94–112 and 118–132 in the CCD of IN (Fig. 5b), and IN 242–260, situated in the C-terminal DNA binding domain (Fig. 5f). Both short peptides, IN1 1–10 and IN1 11–20, bound only two of the four IN1-binding regions (Fig. 5c and d, respectively): IN1 1–10 and IN1 11–20 both bound IN 118–132. In addition, IN1 1–10 bound exclusively to residues 94–112 in the CCD of IN, and IN1 11–20 bound exclusively to residues 242–260 in the DNA binding region of IN.

3. Discussion

To improve the IN1 peptide and to understand its mechanism of interaction with IN, we synthesized several shorter derivatives of the peptide, including its N- and C-terminal halves, and compared their binding and inhibitory properties to those of the parent IN1 peptide. IN1 1–10 and IN1 11–20 bound IN an order of magnitude tighter than the original IN1, whereas the C-terminal peptide IN1 16–20 did not bind IN. This shows that only residues 1–15 of IN1 are important for the interaction with IN. The C3S substitution had only a minor effect on the IN binding and inhibitory activity of IN1 1–10, showing that the cysteine residue itself is not required for IN1 activity and the substitution to serine is well tolerated.

The IN1 sequence is very hydrophobic, especially in the middle region. This led to problems with the synthesis of several other derivatives of the original IN1. IN1 11–20 was optimally synthesized using the ChemMatrix resin.[24–26] The hydrophilic nature of this resin minimizes aggregation of the growing peptide chain, permitting solid phase synthesis of long or difficult peptides.[24]

A striking difference between the IN-binding properties of the long and short IN1 peptides is their ability to bind different IN oligomeric states. The long peptide bound preferentially to the IN tetramer while the short peptides bound the IN dimer. Mapping the peptide binding regions on the structure of the IN CCD[28] (PDB#: 2b4j) and DNA binding domain[13] (PDB#: 1ihv) (Fig. 5) revealed the following interactions: IN 118–132 bound to IN1 and to both the N- and C-terminal halves, IN 94–112 bound IN1 as well as its N-terminal half and IN 242–260 bound IN1 as well as its C-terminal half. IN 118–132, which bound the three peptides, is a surface-exposed bended helix (Fig. 5e) and is known to form interactions with other proteins such as the IN cellular cofactor LEDGF/p75[5] and its derived peptides.[29] The peptide binding regions in the CCD of IN are situated at or near the dimerization interface of IN (Fig. 5). Residues 242–260, which are located in the C-terminal DNA binding domain of the protein, are part of a β-sheet (Fig. 5f) and are surface exposed as well. These residues are at the dimerization interface of the DNA binding domain of IN according to the crystal structure.[13] In summary, our results show that the IN1 peptides bind IN in or near regions involved in dimerization, and this may explain the ability of the peptides to affect IN oligomerization. We propose that the full-length IN1

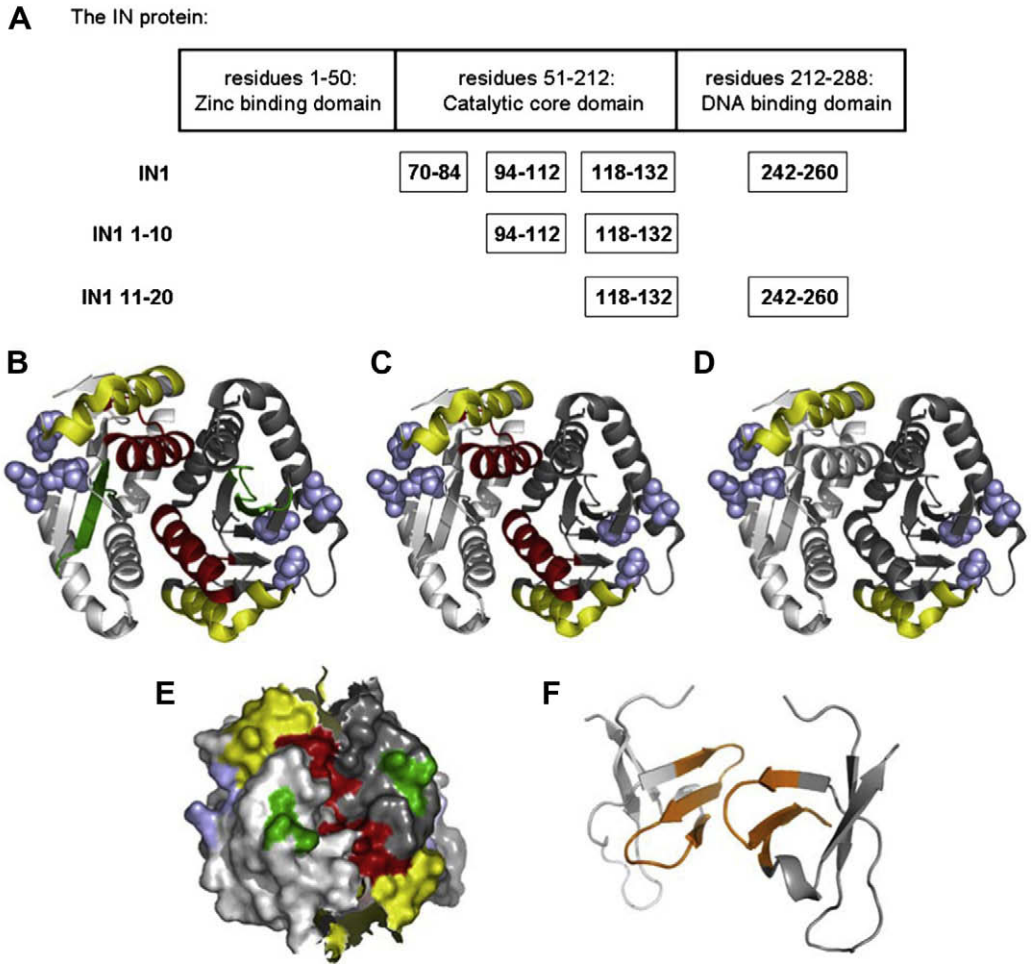

Figure 5. The IN1-derived peptides have different binding regions in IN: (A) schematic presentation of the peptides binding regions in IN; (B–D) The binding sites of (B) IN1, (C) IN1 1–10 and (D) IN1 11–20 on the catalytic core domain of IN[28] (PDB#: 2b4j); (E) a space filling model showing whether the peptide binding regions are surface exposed. In all the structures, the two IN dimers are colored in two shades of blue, IN 70–84 in green, IN 94–108 in red and IN 118–132 in purple. The catalytic residues (D64, D116 and E152) are shown in pink on each monomer; (F) in addition to these sites, both IN1 and IN1 11–20 bind an additional site, IN 242–260 (orange), in the DNA binding domain of IN[13] (PDB#: 1ihv).

peptide may act as a bridge between two dimers, stabilizing a tetramer of IN by forming a dimer of dimers. Both short 10-mer peptides are too short to bridge two dimers together and form a tetramer. However, peptide length is not the only factor that determines the IN oligomeric state to which it binds. LEDGF/p75 361–370, another 10 residue IN-inhibitory peptide found in our labs, bound preferentially to the IN tetramer.[18] Comparing the sequences of IN1 1–10 and LEDGF/p75 361–370 revealed that the sequence of LEDGF/p75 361–370 (WNSLKIDNLDV) is polar, with a slightly negative charge, whereas the sequences of IN1 and its derivatives are highly hydrophobic (Table 1). These hydrophobic properties of the 10-mer IN1 derivatives may be the key for binding the IN dimer, while the polar character of the LEDGF/p75-derived peptide enables it to bind the tetramer. Structural studies need to be performed to gain more insight into the precise mode of the peptides binding to IN.

IN1 and IN1 1–10 have similar IN inhibitory properties, despite the fact that they bind different oligomeric states of IN. Thus, tetramer stabilization[18,19] is not the sole factor responsible for IN inhibition by the IN1 peptide. The binding sites of the IN1 peptides do not overlap the catalytic site of IN so inhibition is also not likely to be caused by masking this site. However, since the peptides inhibited the IN-DNA interaction, and they bind IN at the LEDGF/p75–binding interface,[28] it is possible that inhibition by the IN1 peptides is due to masking the IN interaction with DNA and/or other cofactors, both in the dimeric and tetrameric states.

IN1 1–10 bound IN tighter than the original IN1, and inhibited IN–DNA binding better than IN1. Being shorter, less hydrophobic, more soluble and easier to synthesize than the original IN1 peptide, IN1 1–10 and its C3S derivative are improved leads for further development as anti-HIV lead peptides and as candidates for conversion into non-peptides and small molecules. The information gathered here about the IN1 1–10 binding sites to IN may serve as a basis for such studies.

4. Experimental

4.1. Protein expression and purification

The full-length histidine-tagged IN expression vector was a generous gift from Dr. A. Engelman (Harvard Medicinal School, Boston, MA) and its expression and purification were performed essentially as described.[30]

4.2. Peptide synthesis, labeling and purification

Peptides were synthesized on Rink amide MBHA resin on an ABI 433A peptide synthesizer (Applied Biosystems, Foster City, CA)

using standard Fmoc (9-fluorenylmethoxycarbonyl) chemistry.[31] Trp was added at the peptides N termini for UV spectroscopy. The peptides were labeled at their N termini using 5- and 6-carboxyfluorescein (NOVAbiochem).[32] The peptides were purified on a Gilson HPLC (high pressure liquid chromatography, Middelton, WI) using a reverse-phase C8 semi-preparative column (ACE) with a gradient from 5% to 80% acetonitrile in water (both containing 0.001% (v/v) trifluoroacetic acid). Purity of each peptide was calculated by peak integration after running the sample on a reverse-phase C8 analytical column (ACE) (Table 1b and Fig. S1). The peptides were identified using a Voyager DE-Pro MALDI-TOF mass spectrometer (Applied Biosystems) and NMR (Fig. S2). Peptide concentrations were determined using a UV spectrophotometer (Shimadzu, Kyoto, Japan) as described.[18] Peptide identities were further confirmed by sequential analysis (IN1, IN1 1–10, IN1 1–10 C3S and IN1 16–20) or all residues were accounted for not-sequentially (IN1 11–20) (see Table 1c for assignment and Fig. S2 for spectra).

Additional synthetic methods were tested in order to synthesize the more hydrophobic IN1-derived peptides. Initially, synthesis using a pseudoproline building block (Fmoc-Leu-Thr (Psi Me, Me Pro)-OH (NovaBiochem) in position 5–6 of IN1 11–20) was performed as described,[22,23] but did not result in improved yields. Next, we performed the synthesis on a ChemMatrix (CM) resin as described.[24–26] Briefly, the CM resin was washed twice with methanol, DMF (dimethyl formamide), DCM (dichloromethane), TFA (trifluoracetic acid)/DCM (1:99), DIEA (diisopropylethylamine)/DCM (1:99) and again with DCM. The Fmoc protecting group was removed using the standard procedure of resin incubation for 30 min with 20% piperidine in NMP (*N*-methylpyrrolidone) at room temperature, followed by washes with NMP. 0.1 mmol of the C-terminal amino acid (Fmoc-Arg(Pbf)-OH) was pre-activated for 10 min with HOAt (1-hydroxy-7-azabenzotriazole), HATU (2-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate methanaminium) and DIEA in DMF. The pre-activated amino acid was added to the CM resin and incubated overnight at room temperature. The remaining amino acids were coupled on the ABI 433A peptide synthesizer as described before. In order to obtain improved yields of IN1, we synthesized it on a Liberty Microwave-Assisted Peptide Synthesizer (CEM) using standard Fmoc chemistry.

4.3. Fluorescence anisotropy

Measurements were performed at 10 °C using a PerkinElmer (Waltham, MA) LS-55 luminescence spectrofluorimeter equipped with a Hamilton Microlab 500 dispenser.[33,34] The fluorescein-labeled peptide or DNA (100 nM, 10 nM, respectively, in 1 ml 20 mM Tris buffer, pH 7.4, 185 mM NaCl, 5% glycerol) was placed in the cuvette and the non-labeled IN protein (200 μl, 10 μM) was added in aliquots of 5 μl at 90 s intervals. The total fluorescence and anisotropy were measured after each addition by using an excitation wavelength of 480 nm and an emission wavelength of 530 nm. Data were fit to the Hill equation.

$$R = R_0 + \frac{\Delta R * (K_a^n * [\mathrm{IN}]^n)}{1 + K_a^n * [\mathrm{IN}]^n}$$

(R = measured anisotropy, ΔR = the amplitude of the anisotropy change from R_0 (free peptide) to peptide in complex, [IN] = the added concentration of IN and K_a = the association constant).

For the experiments testing the peptides effect on the IN–DNA binding, the non-labeled IN protein was preincubated for 20 min with the non-labeled IN1 and IN1-derived peptides (500 nM), and then titrated into the fluorescein-labeled LTR DNA solution containing the same peptide concentration (10 nM) in 20 aliquots of 10 μl at 1 min intervals.

4.4. Analytical gel filtration

Analytical gel filtration of IN (10 μM) was performed on an AKTA design Explorer instrument equipped with a Monitor UV-900 detector and the Unicorn software package using a Superose 12 analytical column 30 × 1 cm (GE Healthcare-Pharmacia Corp.) equilibrated with a buffer (20 mM tris, pH 7.4, 1 M NaCl and 10% glycerol), as described.[17–19] Proteins were eluted with a flow rate of 0.5 ml/min at 4 °C and the elution profile was monitored by UV absorbance at 220 nm. The column was calibrated with molecular weight standards (GE Healthcare-Amersham Pharmacia).

4.5. IN integration activity assay

Determination of the IN enzymatic activity by a quantitative assay was performed as described.[17] When peptide inhibition was studied, the IN (390 nM) was incubated with the peptide at the indicated molar ratio and the DNA substrate.

4.6. Cell penetration experiments

The fluorescein-labeled peptides (10 μM in PBS (phosphate buffered saline)) were incubated with HeLa cells for 2 h at 37 °C. After three washes with PBS, the cells were visualized by confocal microscopy.

4.7. MAGI

The MAGI assay was used as described,[27] to quantify the inhibition of HIV-1 infectivity by the peptides. TZM-bl cells were grown in 96-well plates at 10×10^3 cells per well; peptides were added after 12 h of incubation at 37 °C, and after an additional 2 h of incubation, the cells were infected with 50 μl of serially diluted HIV-1 Δenv/VSV-G virus as previously described. The Δenv/VSV-G virus was used since for the MAGI assay only one cycle of infection is required. Two days post-infection, cultured cells were fixed and β-galactosidase was estimated exactly as described previously. Blue cells were counted under a light microscope at a magnification of ×200.

4.8. Infection of cultured lymphocyte cells with HIV-1

Cultured lymphocytes (1×10^5) were centrifuged for 5 min at 2000 rpm and after removal of the supernatant, the cells were resuspended in 0.2–0.5 ml of RPMI 1640 medium containing virus at a multiplicity of infection (MOI) of 0.01 and 5. Following absorption for 2 h at 37 °C, the cells were washed to remove unbound virus and then incubated at the same temperature for an additional 1–8 days.

4.9. PCR analysis of early genes

The viral DNA of the early genes of HIV-1 was estimated as described previously.[20]

4.10. Quantitative analysis of the copy numbers of HIV-1 DNA integrated into the cellular genome

The integration reaction in cultured cells was monitored as follows: Sup T1 cells were incubated for 2 h with the peptides, followed by infection with a HIV-1 Δenv/VSV-G virus at MOI of 2 (as described above) for 24 h. PCR was performed exactly as described previously.[20]

4.11. ELISA-based binding assay

Peptide–peptide binding was estimated using an ELISA-based binding assay as described.[17] A library of 15 residue peptides covering the full length of the IN protein was obtained from the National Institute of Allergy and Infectious Diseases, NIH.

4.12. NMR assignment

Peptide samples IN1 (0.5 mM in 10% D_2O/H_2O, pH 4.2), IN1 1–10 (0.5 mM in 10% D_2O/H_2O, pH 4.5), IN1 11–20 (d_6-DMSO), IN1 1–10 C3S (1.0 mM in 10% D_2O/H_2O, pH 3.3), and IN1 16–20 (1.8 mM in 10% D_2O/H_2O, pH 4.1) were dissolved from lyophilized form. NMR experiments were performed on a Bruker Avance 600 MHz DMX spectrometer at 25 °C. The transmitter frequency was set on the HDO signal, which was calibrated at 4.773 ppm or in the center of the DMSO-d_6 spectrum with the solvent peak used for calibration at 2.50 ppm. TOSCY, using the MLEV-17 pulse scheme for the spin lock[35] and NOESY[36,37] experiments were acquired using the Watergate W5 pulse sequence with gradients and double echo for water suppression.[35,38] Spectra were processed and analyzed with TopSpin (Bruker Analytische Messtechnik GmbH). Resonance assignment followed the sequential assignment methodology developed by Wüthrich.[39]

Table 1c presents assignments for the IN1 peptide and derivatives (see Fig, S2 for spectra). IN1 and IN1 1–10 C3S peptides were unambiguously assigned according to their TOCSY and NOESY spectra. IN1 1–10 was assigned based on the assignment of its mutation IN1 1–10 C3S, where all resonances were in good correlation apart from the mutated residues which gave unique patterns and a shift in Leu4, adjacent to the mutation. IN1 16–20 gave a unique spectrum and was assigned according to its TOCSY spectrum. IN1 11–20 was sparingly soluble in water, so its TOCSY spectrum was acquired in DMSO-d_6 and all residues were present in the expected count.

Acknowledgements

This study was supported by a starting grant from the European Research Council (ERC) (for AF), by a Nophar grant from the Chief Scientist, Ministry of Industry, Israel, with Johnson and Johnson (for AF and A. Loyter) and by the Israel Science Foundation (for A. Loyter). We thank Dr. Mario Lebendiker for his help with the gel filtration experiments.

The following reagent was obtained through the AIDS Research and Reference Reagent Program of the Division of AIDS, National Institute of Allergy and Infectious Diseases, NIH: HIV-1 Consensus B Pol (15-mer) peptides (complete set from National Institute of Allergy and Infectious Diseases Division of AIDS).

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2009.09.053.

References and notes

1. Craigie, R. *J. Biol. Chem.* **2001**, *276*, 23213.
2. Hazuda, D. J.; Young, S. D.; Guare, J. P.; Anthony, N. J.; Gomez, R. P.; Wai, J. S.; Vacca, J. P.; Handt, L.; Motzel, S. L.; Klein, H. J.; Dornadula, G.; Danovich, R. M.; Witmer, M. V.; Wilson, K. A.; Tussey, L.; Schleif, W. A.; Gabryelski, L. S.; Jin, L.; Miller, M. D.; Casimiro, D. R.; Emini, E. A.; Shiver, J. W. *Science* **2004**, *305*, 528.
3. Lataillade, M.; Kozal, M. J. *AIDS Patient Care STDS* **2006**, *20*, 489.
4. Pommier, Y.; Johnson, A. A.; Marchand, C. *Nat. Rev. Drug Disc.* **2005**, *4*, 236.
5. Engelman, A.; Mizuuchi, K.; Craigie, R. *Cell* **1991**, *67*, 1211.
6. Guiot, E.; Carayon, K.; Delelis, O.; Simon, F.; Tauc, P.; Zubin, E.; Gottikh, M.; Mouscadet, J. F.; Brochon, J. C.; Deprez, E. *J. Biol. Chem.* **2006**, *281*, 22707.
7. Chen, A.; Weber, I. T.; Harrison, R. W.; Leis, J. *J. Biol. Chem.* **2006**, *281*, 4173.
8. Deprez, E.; Tauc, P.; Leh, H.; Mouscadet, J. F.; Auclair, C.; Brochon, J. C. *Biochemistry* **2000**, *39*, 9275.
9. Deprez, E.; Tauc, P.; Leh, H.; Mouscadet, J. F.; Auclair, C.; Hawkins, M. E.; Brochon, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10090.
10. Faure, A.; Calmels, C.; Desjobert, C.; Castroviejo, M.; Caumont-Sarcos, A.; Tarrago-Litvak, L.; Litvak, S.; Parissi, V. *Nucleic Acids Res.* **2005**, *33*, 977.
11. Cai, M.; Zheng, R.; Caffrey, M.; Craigie, R.; Clore, G. M.; Gronenborn, A. M. *Nat. Struct. Biol.* **1997**, *4*, 567.
12. Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Engelman, A.; Craigie, R.; Davies, D. R. *Science* **1994**, *266*, 1981.
13. Lodi, P. J.; Ernst, J. A.; Kuszewski, J.; Hickman, A. B.; Engelman, A.; Craigie, R.; Clore, G. M.; Gronenborn, A. M. *Biochemistry* **1995**, *34*, 9826.
14. de Soultrait, V. R.; Caumont, A.; Parissi, V.; Morellet, N.; Ventura, M.; Lenoir, C.; Litvak, S.; Fournier, M.; Roques, B. *J. Mol. Biol.* **2002**, *318*, 45.
15. Sourgen, F.; Maroun, R. G.; Frere, V.; Bouziane, M.; Auclair, C.; Troalen, F.; Fermandjian, S. *Eur. J. Biochem.* **1996**, *240*, 765.
16. Zhao, L.; O'Reilly, M. K.; Shultz, M. D.; Chmielewski, J. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1175.
17. Armon-Omer, A.; Levin, A.; Hayouka, Z.; Butz, K.; Hoppe-Seyler, F.; Loya, S.; Hizi, A.; Friedler, A.; Loyter, A. *J. Mol. Biol.* **2008**, *376*, 971.
18. Hayouka, Z.; Rosenbluh, J.; Levin, A.; Loya, S.; Lebendiker, M.; Veprintsev, D.; Kotler, M.; Hizi, A.; Loyter, A.; Friedler, A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8316.
19. Hayouka, Z.; Rosenbluh, J.; Levin, A.; Maes, M.; Loyter, A.; Friedler, A. *Biopolymers* **2008**, *90*, 481.
20. Rosenbluh, J.; Hayouka, Z.; Loya, S.; Levin, A.; Armon-Omer, A.; Britan, E.; Hizi, A.; Kotler, M.; Friedler, A.; Loyter, A. *J. Biol. Chem.* **2007**, *282*, 15743.
21. Hopkins, A. L.; Groom, C. R. *Nat. Rev. Drug Disc.* **2002**, *1*, 727.
22. Keller, M.; Miller, A. D. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 857.
23. Schmiedeberg, N.; Kessler, H. *Org. Lett.* **2002**, *4*, 59.
24. Garcia-Martin, F.; Quintanar-Audelo, M.; Garcia-Ramos, Y.; Cruz, L. J.; Gravel, C.; Furic, R.; Cote, S.; Tulla-Puche, J.; Albericio, F. *J. Comb. Chem.* **2006**, *8*, 213.
25. Garcia-Martin, F.; White, P.; Steinauer, R.; Cote, S.; Tulla-Puche, J.; Albericio, F. *Biopolymers* **2006**, *84*, 566.
26. Marani, M. M.; Paradis-Bas, M.; Tulla-Puche, J.; Cote, S.; Camperi, S. A.; Cascone, O.; Albericio, F. *J. Comb. Chem.* **2007**, *9*, 395.
27. Kimpton, J.; Emerman, M. *J. Virol.* **1992**, *66*, 2232.
28. Cherepanov, P.; Ambrosio, A. L.; Rahman, S.; Ellenberger, T.; Engelman, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17308.
29. Levin, A.; Hayouka, Z.; Helfer, M.; Brack-Werner, R.; Friedler, A.; Loyter, A. *PLoS ONE* **2009**, *4*, e4155.
30. Jenkins, T. M.; Engelman, A.; Ghirlando, R.; Craigie, R. *J. Biol. Chem.* **1996**, *271*, 7712.
31. Chang, C. D.; Meienhofer, J. *Int. J. Pept. Protein Res.* **1978**, *11*, 246.
32. Weber, P. J.; Bader, J. E.; Folkers, G.; Beck-Sickinger, A. G. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 597.
33. Friedler, A.; Hansson, L. O.; Veprintsev, D. B.; Freund, S. M.; Rippin, T. M.; Nikolova, P. V.; Proctor, M. R.; Rudiger, S.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 937.
34. Friedler, A.; Veprintsev, D. B.; Rutherford, T.; von Glos, K. I.; Fersht, A. R. *J. Biol. Chem.* **2005**, *280*, 8051.
35. Bax, A. *J. Magn. Reson.* **1985**, *65*, 142.
36. Kumar, A.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *96*, 1156.
37. Macura, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 95.
38. Liu, M. L.; Mao, X. A.; Ye, C. H.; Huang, H.; Nicholson, J. K.; Lindon, J. C. *J. Magn. Reson.* **1998**, *132*, 125.
39. Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley and Sons, New York, 1986.